Blood RNA Sequencing Confirms Upregulated *BATF2* and *FCGR1A* Expression in Children with Autism Spectrum Disorder

Irena Voinsky  
Tel Aviv University

Yazeed Zoabi  
Tel Aviv University

Noam Shomron  
Tel Aviv University

Moria Harel  
Shaare Zedek Medical Center

Hanoch Cassuto  
Shaare Zedek Medical Center

Joseph Tam  
The Hebrew University of Jerusalem

Sirish Bennuri  
University of Arkansas for Medical Sciences and Arkansas Children's Research Institute

Shannon Rose  
University of Arkansas for Medical Sciences and Arkansas Children's Research Institute

Richard E. Frye  
University of Arizona College of Medicine-Phoenix

Adi Aran  
Shaare Zedek Medical Center

David Gurwitz (✉ gurwitz@tauex.tau.ac.il)  
Tel Aviv University

Research Article

**Keywords:** autism spectrum disorder, blood, peripheral blood mononuclear cells, RNA sequencing, real-time PCR, cancer

**Posted Date:** December 13th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-1113898/v1
Abstract

Mutations in over 100 genes are implicated in autism spectrum disorder (ASD). DNA mutations and epigenomic modifications also contribute to ASD. Transcriptomics analysis of blood samples may offer clues for pathways dysregulated in ASD. To expand and validate published findings of RNA-sequencing (RNA-seq) studies, we performed RNA-seq of whole blood samples from a discovery cohort of eight children with ASD compared with nine age- and sex-matched neurotypical children. This revealed 10 genes with differential expression. Using real-time PCR, we compared whole blood samples from 35 children with ASD and 21 matched neurotypical children for the 10 dysregulated genes detected by RNA-seq. This revealed higher expression levels of the proinflammatory transcripts \textit{BATF2} and \textit{FCGR1A}, and lower expression levels of the anti-inflammatory transcripts \textit{ISG15} and \textit{MT2A} in the ASD compared to the control group. \textit{BATF2} and \textit{FCGR1A} were recently reported as upregulated in blood samples of Japanese adults with ASD. Coupled with that publication, our findings support involvement of these genes in ASD phenotypes, independent of age and ethnicity. Upregulation of \textit{BATF2} and \textit{FCGR1A} and downregulation of \textit{ISG15} and \textit{MT2A} were reported to reduce cancer risk. Implications of the dysregulated genes for proinflammatory phenotypes, immunity, and cancer risk in ASD are discussed.

Introduction

Autism spectrum disorder (ASD) is the most complex and heterogenous human neurological disorder. Inherited or de novo mutations in over 100 known genes are already implicated in ASD, yet, most incidences remain unexplained. ASD may result from detrimental epigenetic modifications during early embryonic development\textsuperscript{1,2}. Further contributors to the highly diverse ASD phenotypes include unique combinations of common variants in many genes (a high polygenic risk score); prenatal environmental influences\textsuperscript{3}; mitochondrial deficiencies\textsuperscript{4,5}; a chronic pro-inflammatory state\textsuperscript{6,7}; psychiatric and neurologic comorbidities\textsuperscript{8}; and aberrant gut microbiome composition\textsuperscript{9}. Together, the diverse phenotypes contribute to difficulties in early diagnosis of autism among children, which is crucial for early treatment and parental guidance\textsuperscript{10,11}.

The involvement of epigenetics in ASD phenotypes hinders efforts to establish improved diagnostic tools for early identification of ASD and its sub-phenotypes, using genomic, biochemical, and metabolomics analyses\textsuperscript{12}. Brain imaging diagnostic tools such as fMRI have been also suggested; however, such tools still have low stratification value and low reproducibility, and thus require further studies before their incorporation in routine pediatric neurology practice\textsuperscript{5,13,14}. Genome-wide transcriptomic studies examine changes in gene expression that also reflect epigenetic DNA modifications, thereby circumventing complex and costly technologies such as DNA methylome profiling\textsuperscript{15}. Indeed, since the completion of the human genome project, numerous studies have applied genome-wide transcriptomics, initially using microarrays and more recently using RNA-sequencing (RNA-seq) technologies, for identifying altered gene expression patterns in blood samples from individuals with ASD compared to neurotypical controls.
Venous blood samples represent an accessible, affordable, and readily available biological resource for establishing differential diagnosis, prognosis, and subtyping of complex disorders such as ASD. Indeed, the majority of transcriptomic studies in humans have utilized venous blood samples. Hence, we aimed to perform comparative RNA-seq in an independent cohort of whole blood samples from children with ASD and a neurotypical control group, to identify the top dysregulated gene transcripts, and to assess (by literature survey) if some of the detected dysregulated genes were already reported as dysregulated in ASD. Here we report that our comparative RNA-seq identified two upregulated mRNA transcripts, \textit{BATF2} and \textit{FCGR1A}, in whole blood samples of Israeli children with ASD compared to a control group of neurotypical children. These genes were recently reported as upregulated in whole blood from Japanese adults with ASD.

\section*{Results}

\subsection*{RNA sequencing and real-time PCR validation}

Whole blood samples were collected from 36 children with ASD and 21 NT children at Shaare Zedek Medical Center (Jerusalem, Israel). PBMCs were collected from 40 children with ASD and 26 NT children at the ACH (Little Rock, AR, USA.), as described in the Methods section. The demographics of the Israeli and U.S. cohorts are presented in Table 1. RNA was extracted from these blood samples. RNA from eight males with ASD (mean age 13.5±2.6 y) and nine NT males (controls; mean age 15.4±1.9 y) were applied for RNA-seq as described in the Methods. Bioinformatics analysis of the RNA-seq reads, followed by p-value adjustment for genome-wide transcriptomics, identified 10 dysregulated genes with differential expression in whole blood in the ASD compared to the control group, \( p_{\text{adj}}<0.05 \) (Table 2). Next, we performed real-time PCR validation for these dysregulated genes in our entire cohort of whole blood samples (36 ASD and 21 matched NT Israeli children; Table 1). Our findings showed that four of the genes detected as dysregulated by RNA-seq were confirmed as dysregulated in the entire Israeli cohort of whole blood RNA samples (Fig. 1). We observed upregulated expression of \textit{BATF2} and \textit{FCGR1A} (FD=2.03, \( p=0.004; \text{FD}=1.5, p=0.0013 \)) and downregulated expression of \textit{ISG15} and \textit{MT2A} (FD=0.64, \( p=0.0074; \text{FD}=0.74, p=0.013 \)) in the ASD compared to the control group. Other genes found as dysregulated in RNA-seq from whole blood of the two groups (Table 2) were not validated in the entire Israeli cohort; their real-time PCR findings are shown in \textbf{Supplementary Fig. S1}.

Next, we assessed RNA levels of the 10 dysregulated genes detected by our RNA-seq of whole blood (Table 2) in RNA extracted from PBMCs from a second cohort of children with ASD and NT controls (U.S. cohort; Table 1). None of these genes showed differential expression in PBMC samples from the ASD versus the control group. The NT controls of the U.S. cohort included both NT siblings of the ASD group and unrelated NT controls (Table 1). We therefore compared the RNA expression levels in PBMCs from children with ASD, separately to those of their NT siblings and to those of unrelated NT controls, for the same 10 genes. The comparison to NT siblings indicated upregulated \textit{SERPING1} (FD=1.90; \( p=0.023 \)) in
the children with ASD. The real-time PCR findings from the ASD and control PBMC samples are shown in Supplementary Fig. S2 and Supplementary Table S2.

**Correlations of whole blood gene expression levels with behavioral scores**

We looked for possible correlations of the whole blood RNA levels of the genes detected as dysregulated in our RNA-seq, with behavioral phenotypes of the same children with ASD. We observed negative correlations between whole blood RNA levels of *BATF2* or *SERPING1* and the scores of the individual teacher SRS (tSRS). We observed positive correlations of RNA levels of *LY6E* with VABS socialization domain scores and VABS Composite scores, and of RNA levels of *ISG15* with cbcl scores (Fig. 2). We also observed positive correlations of PBMC RNA levels of *BATF2*, *MT2A*, *LY6E*, and *ISG15* with Aberrant Behavior Checklist (ABC) scores (Fig. 3).

**Correlations of whole blood gene expression levels with serum endocannabinoids**

Reduced levels of several endocannabinoids were reported in serum samples of children with ASD compared with controls\(^\text{16}\). These serum samples were from the same Israeli children from whom we collected whole blood samples and analyzed the RNA samples used in the current study. Therefore, in each of the study participants (both the ASD and control groups), we looked for correlations between blood RNA expression levels of the top dysregulated genes (Table 2) and each of the serum endocannabinoid levels reported by Aran et al. 2019\(^\text{16}\). The significant correlations are shown in Fig. 4. Further correlations for whole blood mRNA expression and serum endocannabinoid levels are listed in Supplementary Table S3. The strongest correlation observed was a negative correlation between whole blood *LY6E* mRNA expression levels and serum N-palmitoylethanolamine (PEA) in the NT children (PEA; \(r=0.7298, p=0.0004\)), while no such correlation was observed for the ASD group (Supplementary Fig. S3). Further correlations between blood mRNA expression levels and serum endocannabinoid levels were detected by combining the ASD and control groups for each correlation plot (Supplementary Fig. S4).

**Discussion**

**Upregulated expression of BATF2 and FCGR1A in Israeli children with ASD confirms findings in Japanese adults with ASD**

The upregulated expression levels of both *BATF2* and *FCGR1A* in our Israeli cohort of whole blood samples from children with ASD (Fig. 1) corroborate findings from a recent Japanese RNA-seq study in whole blood from adults with ASD\(^\text{17}\). To our knowledge our study provides the first validation of a genome-wide RNA-seq study with ASD blood samples. Our literature search of RNA-seq studies (Table 3; see Methods) revealed large variation in findings from earlier transcriptomic studies (using either RNA-
seq or RNA microarray technologies) in whole blood samples of individuals with ASD compared to matched NT controls. Notably, each of the dysregulated genes listed in our literature survey was mentioned in only a single study (or a single meta-analysis in regard to the meta-analysis by Lee et al., 2019\textsuperscript{18}). Therefore, our current study appears to be the first validation of any dysregulated gene in blood samples from individuals diagnosed with ASD. Moreover, our validation was done on children with ASD (Table 1), in contrast to adults with ASD in the mentioned Japanese study\textsuperscript{17}.

Our RNA-seq detected \textit{SERPING1} as the top upregulated gene in whole blood from children with ASD (FD=3.4962; $P_{\text{adj}}=0.0072$; Table 2). \textit{SERPING1} was among the upregulated genes in whole blood from Japanese adults with ASD\textsuperscript{17}. Yet, our real-time PCR experiments could not validate this finding in the entire whole blood samples (\textit{Supplementary Fig. S1}). Likewise, \textit{SERPING1} expression was similar in PBMC-derived RNA samples from children with ASD and NT children (\textit{Supplementary Table S2}). Nonetheless, our real-time PCR experiments in PBMC-derived RNA samples (our U.S. cohort) indicated upregulated expression of \textit{SERPING1} in children with ASD compared with their NT siblings (\textit{Supplementary Table S2}). These findings exemplify an advantage of including NT siblings of children with ASD in autism research studies.

Our findings thus suggest a central role in ASD for the upregulated genes identified, independent of age and ethnicity. Further studies, including transcriptomic studies with brain tissues from ASD animal models, are needed to elucidate the relevance of the dysregulated genes for ASD behavioral scores (Fig. 2 & Fig. 3) and their implications for ASD phenotypes.

### Dysregulated ASD genes and cancer

All four genes that were found in our study to present dysregulated mRNA transcript levels in blood from children with ASD have been investigated mostly in the context of cancer. Both \textit{BATF2} and \textit{FCGR1A}, detected in our study as upregulated in blood samples of children with ASD, code for cancer protective proteins. \textit{BATF2} was shown to have an antitumor effect in a mouse model through upregulation of IL-12 p40 in tumor-associated macrophages, leading to CD8$^+$ T-cell activation and tumor accumulation\textsuperscript{19}. Among other cancers, \textit{BATF2} was demonstrated as a tumor suppressor of gastric cancer\textsuperscript{20}, glioblastoma\textsuperscript{21}, and esophageal squamous cell carcinoma\textsuperscript{22}. Higher tumor \textit{FCGR1A} expression correlated with improved prognosis in laryngeal carcinoma\textsuperscript{23}, as well as in cervical cancer and melanoma, in which it was associated with increased tumor infiltration of CD4$^+$ and CD8$^+$ T cells and dendritic cells\textsuperscript{24}.

Additionally, higher expression levels of both \textit{ISG15} and \textit{MT2A}, found in this study as downregulated in blood samples from children with ASD, were reported to be associated with worse cancer prognosis. The protein coded by \textit{ISG15} (interferon-stimulated gene 15 ubiquitin like modifier) was implicated in autophagy, exosome secretion, DNA repair, and immune modulation pathways; it is also a known tumor promoter by suppressing immune cell tumor infiltration\textsuperscript{25}. \textit{ISG15} was shown to drive tumorigenesis and metabolic plasticity of pancreatic cancer, suggesting that its inhibition may be a treatment option for
pancreatic cancer. Higher ISG15 expression was also associated with poor prognosis in breast cancer. The protein coded by MT2A, metallothionein 2A, is the major metallothionein in humans, and serves as a chelator of intracellular zinc ions and protects cells against free radicals. MT2A is upregulated in most cancers, and contributes to their chemotherapy resistance by chelation of zinc and platinum-containing drugs and by its action on p53 zinc-dependent activity. MT2A upregulation results in p53 misfolding secondary to zinc chelation, while low cellular MT2A levels allow proper p53 function as a genome stability guardian. Lastly, downregulated C1 Inhibitor (encoded by SERPING1) was shown to increase cancer risk. Hence, the upregulated SERPING1 observed in sub-cohorts of this study may also contribute to reduced cancer risk in children with ASD.

Taken together, the upregulation of both BATF2 and FCGR1A, and the downregulation of both ISG15 and MT2A, as we detected in blood samples from children with ASD, all suggest a reduced risk of cancer. Indeed, a huge reduction in cancer risk (OR=0.06; 95% CI: 0.02, 0.19; p<0.0001) was reported among children with ASD aged 0 to 14 years compared with matched controls. These authors compared cancer rates in 1,837 individuals with ASD and in 9,336 controls in the registry of the University of Iowa Hospitals and Clinics. They observed that the large gap in cancer rates between individuals with ASD and controls was lower at older ages, being only 2-fold less among individuals with ASD aged above 55 years compared with controls.

Our findings on upregulated BATF2 and FCGR1A, and downregulated ISG15 and MT2A (or possibly some of these genes) in children with ASD thus seem to agree with the findings of the Darbro et al.2016 epidemiologic survey. Albeit, we did not identify similarly large studies on cancer risk among children with ASD. The only other epidemiologic study reporting reduced cancer risk among individuals with ASD was smaller (91 individuals with ASD and 6,186 sex- and birth-year controls), and was based on death records, thus on older individuals. For all ages combined, it reported a 4.3-fold reduced risk of death from metastatic cancer compared with controls. However, an earlier study reported 1.95-fold higher cancer incidence among males with ASD based on a Taiwanese cancer registry; the elevated cancer risk was particularly high (3.58-fold) for individuals with ASD aged 15-19 years. Additionally, higher cancer mortality (OR=1.80) among individuals with ASD was reported in a study on premature mortality. Yet, the latter study did not include breakdown of death by age. Hence the cancer risk among individuals with ASD remains controversial. Epidemiologic studies with larger cohorts are required to assess the cancer risk among children with ASD compared with NT children.

Dysregulated ASD genes and immunity

Among the common phenotypic features observed in ASD is innate immune system dysregulation, leading to a chronic pro-inflammatory state. The innate immune pathways affected in ASD include signaling mediated via cytokines, hepatocyte growth factor receptor, microglia, and the complement system. These suggest a role for aberrant immune function in the broad ASD phenotypes. A recent RNA-seq study of whole blood from adults with ASD found dysregulated transcription of genes implicated in innate and adaptive immunity. These included upregulated expression of BATF2 and
FCGR1A\textsuperscript{17}, as confirmed in our current study of whole blood from children with ASD. The consequences to the immune system, of dysregulation in children with ASD of the four genes observed in our study (Fig. 1), is discussed in the above section. Notably, BATF2 was shown to promote inflammation in response to lipopolysaccharides or infection\textsuperscript{38}, while ISG15 is known to promote anti-inflammatory pathways\textsuperscript{39,40}. Thus, the upregulation of \textit{BATF2} mRNA, as well as the downregulation of \textit{ISG15} mRNA observed in our study supports the involvement of the pro-inflammatory phenotypes that have often been observed in ASD\textsuperscript{6,7,36}.

\textit{Lack of validation of the genes that were dysregulated in whole blood RNA-seq, in PBMC samples}

Our real-time PCR experiments did not validate any of the 10 dysregulated genes detected by our RNA-seq of whole blood (Israeli cohort; Table 1) in RNA extracted from PBMCs of a second cohort of children with ASD and NT controls (U.S. cohort; Table 1). Nonetheless, comparing children with ASD to their NT siblings indicated upregulated \textit{SERPING1} in the PBMCs of those with ASD (\textit{Supplementary Table S2}). The real-time PCR findings from the PBMC samples, as presented in \textit{Supplementary Fig. S2} and compared with Fig. 1, suggest that the source of the other dysregulated transcripts detected in our RNA-seq of whole blood RNA (Table 2) mostly represent neutrophil RNA. This is because these cells (which represent the major source of blood-derived RNA) are depleted during isolation of PBMCs from whole blood (neutrophils have higher density than PBMCs, and are removed during PBMC separation, together with erythrocytes). This conclusion is rational considering that neutrophils are the key players in inflammation\textsuperscript{41,42}, and that individuals with ASD often display pro-inflammatory phenotypes\textsuperscript{6,7,36}. Indeed, inflammatory signaling and reactive oxygen species mediators were shown to be upregulated in neutrophils of children with ASD\textsuperscript{43}.

\textit{Correlations of whole blood gene expression levels with serum endocannabinoids}

Recent years have seen growing interest in studying the use of cannabinoid drugs for treating behavioral and social deficits of individuals with ASD\textsuperscript{44}. The endocannabinoid system was reported to be dysregulated in various animal models of ASD\textsuperscript{45}. Children with ASD were reported to have lower serum endocannabinoids\textsuperscript{16}, and rare mutations in endocannabinoid pathway genes were implicated in some persons with ASD\textsuperscript{46}. Circulating endocannabinoids are derived from multiple tissues\textsuperscript{47}. However, plasma endocannabinoid levels were demonstrated to reflect brain concentrations\textsuperscript{48}. Hence, the correlations reported here for blood expression of some of the dysregulated genes with serum endocannabinoids (Fig. 4) can shed light on the pathophysiology of ASD. However, studies in animal models of ASD, which allow measurements of endocannabinoid levels and transcriptomics in brain tissues during different stages of pre- and postnatal development\textsuperscript{49,50}, are required for exploring these correlations. The endocannabinoid PEA was reported to display anti-inflammatory\textsuperscript{13,51}, antiepileptic\textsuperscript{52}, and antineuropathic\textsuperscript{53} properties. The \textit{LY6E} cell surface protein was shown to be upregulated by several inflammatory cytokines, including interferons, TNF-alpha, and IL-1 alpha\textsuperscript{54}. As expected, we found a
strong negative correlation between the anti-inflammatory endocannabinoid PEA and the pro-inflammatory transcript \textit{LY6E} in NT children (\textbf{Supplementary Table S3}, \textit{r}=-0.73, \textit{p}=0.0004). This expected negative correlation was not observed in children with ASD, thus demonstrating another example of dysregulated immunity in ASD. Further studies are needed to clarify the relation of the dysregulated blood transcriptomics to the reduced serum endocannabinoids observed in children with ASD\textsuperscript{16}.

\textbf{Limitations}

Our current study has several limitations. First, the small cohorts do not enable separate analysis for males and females. As all the participants in the RNA-seq discovery cohort and the majority of the participants in the entire Israeli cohort were male, the relevance of our findings to female children with ASD needs to be clarified. Second, the whole blood samples (Israeli cohort) and PBMC samples (U.S. cohort) were collected separately, prior to conducting the RNA-seq project; hence, we did not have both whole blood and freshly separated PBMCs from the same individuals. Since the RNA-seq project, its analysis, and real-time PCR validation experiments were conducted from October 2020 to September 2021, recruitment of the children was affected by the Covid-19 pandemic and related lockdowns. Additionally, the children with ASD in the U.S. cohort were on average four years younger than those of the Israeli cohort (Table 1). Hence, our findings on lack of validation for the whole blood dysregulated genes in PBMC samples from children with ASD require confirmation with a study that compares whole blood and PBMC derived RNAs from the same participants. Lastly, the relevance of our transcriptomic findings for early ASD diagnosis is uncertain, as only a few children in both the Israeli and the U.S. cohorts were under age 4 years, the most crucial period for early ASD diagnosis\textsuperscript{10,11}.

In light of the above mentioned reservations, we conclude that validation in larger cohorts, which will ideally include blood samples from younger children, both males and females, are essential for assessing the potential diagnostic and prognostic values of the dysregulated genes detected in our current study. The implications for the upregulated blood transcription of \textit{BATF2} and \textit{FCGR1A}, and downregulated transcription of \textit{ISG15} and \textit{MT2A}, for the distinctive immune system phenotypes in autism, as well as the controversial published findings on lower cancer risk among children with ASD, merit further studies.

\textbf{Methods}

\textbf{Participants}

The study comprised an Israeli and a U.S. cohort. All experimental protocols were approved by the Shaare Zedek Medical Center Institutional Review Board (Israeli cohort) or the University of Arkansas for Medical Sciences Institutional Review Board (U.S. cohort). The Israeli cohort was described by Aran et al., 2019\textsuperscript{16}. Briefly, children and young adolescents with ASD were recruited as part of a randomized clinical trial (NCT02956226) approved by the Shaare Zedek Medical Center Institutional Review Board and the Israeli Ministry of Health. Unrelated, age- and gender-matched, neurotypical (NT) children who attended regular education and were without any neuropsychiatric diagnosis other than attention deficit hyperactivity...
disorder were recruited through advertisements posted in the surrounding community of Jerusalem, Israel. The U.S. cohort, described by Voinsky et al., 2019⁵⁵, provided peripheral blood mononuclear cell (PBMC) samples. This cohort included children with ASD, their NT siblings without ASD, and an unrelated control group of NT children. The protocol is registered in clinicaltrials.gov as NCT02000284. This part of the study was performed at Arkansas Children’s Research Institute and Arkansas Children’s Hospital (ACH) and approved by the Institutional Review Board at the University of Arkansas for Medical Sciences (Little Rock, AR, USA). For both the Israeli and the U.S. cohorts, the children’s parents provided written informed consent and written assent was obtained from participants when appropriate. The study and its methods were performed in accordance with the approved protocols and the relevant guidelines and regulations.

**Behavioral tests**

Behavioral tests for the participants of the Israeli cohort were described by Aran et al. 2019¹⁶. Briefly, they included scores on the Autism Diagnostic Observation Schedule (ADOS-2), the Vineland Adaptive Behavior Scale Composite and socialization domain scores (VABS), the Social Responsiveness Scale-II (SRS-2), and the Child Behavior Checklist (CBCL). Behavioral tests for the participants of the U.S. cohort included, in addition, scores on the Aberrant Behavior Checklist (ABC)⁵⁵.

**Whole blood and PBMC sample collection**

Whole blood and PBMC samples were collected as described by Aran et al., 2019¹⁶ and Voinsky et al., 2019⁵⁵, respectively. Briefly, venous whole blood non-fasted samples were collected in the morning hours. Whole blood (Israeli cohort) samples were collected into Tempus™ Blood RNA tubes (Applied Biosystems™ Catalog number 4342792, Thermo Fisher Scientific, MA, USA) and tubes were frozen immediately at minus 80°C until the RNA extractions. For PBMC sample preparation (U.S. cohort), whole blood samples were collected in EDTA Vacutainer™ tubes, and PBMCs were separated from fresh blood samples as described⁵⁵ and stored at minus 80°C until the RNA extractions.

**RNA extraction**

RNA was extracted from whole blood samples using the Tempus™ Spin RNA Isolation Kit (Invitrogen™, Thermo Fisher Scientific, MA, USA) by following the manufacturer’s protocol. RNA quality was determined by an automated electrophoresis process using the TapeStation system (Agilent, CA, USA). Samples with high quality RNA were specified as those with RIN (RNA integrity number) >8. Nucleic acid quantitation was carried out by Qubit™ RNA HS Assay Kit (Invitrogen™, Thermo Fisher Scientific, MA, USA). The samples selected for sequencing were diluted to 1 ug/ml.

RNA was extracted from PBMC samples as described previously⁵⁵. Due to low RNA amounts, these samples were not prepared for sequencing, and were applied for assessing the expression of the top transcripts detected by our RNA-seq discovery cohort in PBMC (real-time PCR).

**Library preparation, RNA sequencing, and data processing**
RNA samples were shipped in dry ice to Macrogen Europe BV (Amsterdam, Netherlands) for poly-A mRNA sequencing. Libraries were prepared using TrueSeq stranded total RNA LT sample prep kit (Illumina, CA, USA). RNA sequencing was performed as a paired-end read on an Illumina True-Seq platform. Sequencing depth was ~30 M reads/sample. Raw sequencing data were trimmed using fastp 0.20.0\textsuperscript{56} and aligned to the GRCm38 assembly using STAR 2.6.0c\textsuperscript{57}. DESeq2 1.30.1\textsuperscript{58} and R 3.6 were used for normalization of count data and for differential gene expression analysis.

**Real-time qPCR validation**

Real-time quantitative PCR (qPCR) reactions were performed with 1 µg RNA samples converted to cDNA using qScript cDNA Synthesis Kit (Quanta Bio, MA, USA). Reverse transcription was performed using a thermal cycler over three steps (22°C for 5 min, followed by 42°C for 30 min and 85°C for 5 min). Real-time PCR reactions were performed in mixtures containing 10 ng of cDNA, PerfeCTa SYBR® Green FastMix Kit (Quanta Bio, MA, USA), and Integrated DNA Technologies, Inc. (Leuven, Belgium) primers. \textit{GAPDH} (Glyceraldehyde 3-phosphate dehydrogenase) and \textit{RPLP0} (Ribosomal Protein Lateral Stalk Subunit P0) were used as reference genes for whole blood and PBMC samples, respectively. Forward and reverse primer sequences for RT–qPCR are listed in \textit{Supplementary Table S1}.

**Correlations and statistical analysis**

Real-time qPCR data analysis was conducted using the GraphPad Prism v.9 (San Diego, CA, USA). Normality of the data distribution was evaluated using the Shapiro-Wilk test; continuous variables between two groups were analyzed by the Mann-Whitney test; outliers were detected by the ROUT test. P-values ≤ 0.05 were considered as significant. Behavioral measurements and endocannabinoid measurement data, for correlation analysis, were collected and described in our previous studies\textsuperscript{16}. Correlations were examined by assessing normality of data distribution using the Shapiro-Wilk test, followed by a Spearman correlation test.

**Literature survey of transcriptomic studies in whole blood from individuals with ASD and controls**

We searched the NCBI PubMed database for transcriptomic studies published from 2010 to September 2021, of whole blood samples of individuals with ASD compared to controls. Inclusion criteria were the following words in the abstract or keywords: transcriptomics, RNA-sequencing (or RNA-seq), or microarrays; and also both words autism and blood. RNA-seq studies with other biological human samples (such as PBMCs, lymphoblastoid cell lines, and postmortem tissues), or with fewer than six blood samples, were excluded.

**Declarations**

**Acknowledgments**
The study was supported by a grant from the Israel-United States Science Foundation (BSF grant 2019049) to DG, REF, and SR. This work was carried out as part of fulfillment requirements for the PhD thesis of IV. The authors are grateful for the parents/guardians of the study participants for their consent to the use of their children’s blood samples for research.

Author contributions

I.V. performed the RNA extractions (Israeli cohort) and the real-time PCR experiments (all cohorts), prepared RNA samples for RNA-seq, performed the statistical and correlation analyses and prepared all Figures and Tables. Y.Z. and N.S. analyzed the RNA-seq data. A.A., M.H. and H.C. collected the Israeli cohort. J.T. performed the serum endocannabinoid measurements (Israeli cohort). S.B. performed the RNA extractions (U.S cohort). S.R. and R.E.F. collected the U.S. cohort. D.G., A.A., R.E.F., and S.R. conceived the study and wrote the article. All authors have approved the manuscript.

Competing interests

The authors declare no competing interests or other interests that might be perceived to influence the results and/or discussion reported in this article.

Data availability

The datasets generated during the current study are available from the corresponding authors on reasonable request.

References


Table 1: Israeli and U.S. cohorts of children with autism spectrum disorder (ASD) and neurotypical children (control). Differences in mean age between the groups within each cohort were not significant (p>0.1). However, the mean age and proportion of males were lower in the U.S. than the Israeli cohort. *For the U.S. cohort, the neurotypical control group included both siblings of the ASD group and unrelated children.

<table>
<thead>
<tr>
<th>Israeli cohort (whole blood)</th>
<th>U.S. cohort (PBMC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurotypical controls</td>
<td>Children with ASD</td>
</tr>
<tr>
<td>21</td>
<td>36</td>
</tr>
<tr>
<td>Age (years)</td>
<td>16.3±2.7</td>
</tr>
<tr>
<td>% Male</td>
<td>75%</td>
</tr>
<tr>
<td>Neurotypical controls*</td>
<td>Children with ASD</td>
</tr>
<tr>
<td>26</td>
<td>40</td>
</tr>
<tr>
<td>Age (years)</td>
<td>8.9±4.0</td>
</tr>
<tr>
<td>% Male</td>
<td>76%</td>
</tr>
<tr>
<td></td>
<td>8.4±4.4</td>
</tr>
</tbody>
</table>

**ASD behavioral data**

<table>
<thead>
<tr>
<th></th>
<th>Israeli cohort</th>
<th>U.S. cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>VABS Composite score</td>
<td>N/A</td>
<td>48±13</td>
</tr>
<tr>
<td>VABS socialization</td>
<td>N/A</td>
<td>46±11</td>
</tr>
<tr>
<td>domain score</td>
<td>N/A</td>
<td>66±11</td>
</tr>
<tr>
<td>SRS t score</td>
<td>N/A</td>
<td>78±10</td>
</tr>
<tr>
<td>CBCL score</td>
<td>N/A</td>
<td>55±11</td>
</tr>
<tr>
<td>ABC score</td>
<td>N/A</td>
<td>21±29</td>
</tr>
</tbody>
</table>

VABS: Vineland Adaptive Behavior Scale; SRS: Social Responsiveness Scale; CBCL: Child Behavior Checklist; ABC: Aberrant Behavior Checklist

Table 2: Top transcripts showing differential expression in whole blood RNA sequencing. RNA-seq reads are shown for children with autism spectrum disorder (ASD; N=8) vs. neurotypically developing children (control; N=9) with adjusted p<0.05. The genes are arranged by increasing adjusted p values (P_{adj}). FD, fold difference (ASD vs. controls).
<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene ID</th>
<th>Mean reads</th>
<th>FD</th>
<th>p-value</th>
<th>P_adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>SERPING1</td>
<td>ENSG00000149131</td>
<td>494.495</td>
<td>3.4962</td>
<td>3.73E-07</td>
<td>0.0072</td>
</tr>
<tr>
<td>EFHC2</td>
<td>ENSG00000183690</td>
<td>93.312</td>
<td>0.6086</td>
<td>2.81E-06</td>
<td>0.0261</td>
</tr>
<tr>
<td>BATF2</td>
<td>ENSG00000168062</td>
<td>127.616</td>
<td>3.4360</td>
<td>4.06E-06</td>
<td>0.0261</td>
</tr>
<tr>
<td>CDC20</td>
<td>ENSG00000117399</td>
<td>66.744</td>
<td>3.1642</td>
<td>6.06E-06</td>
<td>0.0292</td>
</tr>
<tr>
<td>FCGR1A</td>
<td>ENSG00000150337</td>
<td>416.310</td>
<td>2.6044</td>
<td>1.37E-05</td>
<td>0.0379</td>
</tr>
<tr>
<td>MT2A</td>
<td>ENSG00000125148</td>
<td>407.810</td>
<td>1.8252</td>
<td>1.35E-05</td>
<td>0.0379</td>
</tr>
<tr>
<td>ISG15</td>
<td>ENSG00000187608</td>
<td>1262.341</td>
<td>3.1128</td>
<td>1.36E-05</td>
<td>0.0379</td>
</tr>
<tr>
<td>FBXO6</td>
<td>ENSG00000116663</td>
<td>365.625</td>
<td>1.5121</td>
<td>1.89E-05</td>
<td>0.0416</td>
</tr>
<tr>
<td>LINC00869</td>
<td>ENSG00000277147</td>
<td>306.080</td>
<td>1.4388</td>
<td>1.94E-05</td>
<td>0.0416</td>
</tr>
<tr>
<td>LY6E</td>
<td>ENSG00000160932</td>
<td>3225.314</td>
<td>1.8098</td>
<td>2.49E-05</td>
<td>0.0481</td>
</tr>
</tbody>
</table>

Due to technical limitations, table 3 is only available as a download in the Supplemental Files section.

Figures
Real-time qPCR validation for whole blood RNA expression levels in ASD and control children (Israeli cohort). Box plots show mean ± SEM mRNA levels for ASD vs. neurotypical control whole blood samples. Outliers were removed and analysis was done using a non-parametric Mann Whitney test. FD and p values are shown for the genes with differential expression in ASD vs. neurotypical controls.
Figure 2

Whole blood mRNA expression levels of BATF2, SERPING1 and LY6E correlate with ASD behavioral scores (Israeli cohort). Correlations for mRNA levels (real-time Pare shown for (a) BATF2 vs. tSRS score (N=27); (b) LY6E vs. VABS socialization domain score (N=32); (c) LY6E vs. VABS Composite score (N=32); (d) ISG15 vs. CBCL score (N=33). The r and p values for each correlation plot (Spearman test) are shown in each panel. See Methods for further information on behavioral scores.
Figure 3

PBMC mRNA expression levels of BATF2, ISG15, MT2A, and LY6E correlate with ASD ABC behavioral scores (U.S. cohort). Correlations for mRNA levels (real-time Pare shown for (a) BATF2 (N=46); (b) MT2A (N=42); (c) LY6E (N=44); (d) ISG15 (N=42). The r and p values for each correlation plot (Spearman test) are shown in each panel. See Methods for further information on ABC scores.

Figure 4

Correlations for whole blood mRNA expression levels with serum endocannabinoid levels in ASD children (Israeli cohort). Correlations are shown for (a) BATF2 and oleoylethanolamide (OEA); (b) ISG15 and arachidonic acid (AA); (c) LINC00869 and Oleoyl serine (OS). The r and p values for each correlation plot (Spearman test) are shown in each panel. Endocannabinoid levels are from Aran et al. 201916. See Methods for further details.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTablesandFiguresVoinskyetalScientificReports25.11.21.docx.pdf
• Table3.jpg